Where to attend next: guiding refreshing of visual, spatial, and verbal representations in working memory

Souza, Alessandra S ; Vergauwe, Evie ; Oberauer, Klaus

Abstract: One of the functions that attention may serve in working memory (WM) is boosting information accessibility, a mechanism known as attentional refreshing. Refreshing is assumed to be a domain-general process operating on visual, spatial, and verbal representations alike. So far, few studies have directly manipulated refreshing of individual WM representations to measure the WM benefits of refreshing. Recently, a guided-refreshing method was developed, which consists of presenting cues during the retention interval of a WM task to instruct people to refresh (i.e., attend to) the cued items. Using a continuous-color reconstruction task, previous studies demonstrated that the error in reporting a color varies linearly with the frequency with which it was refreshed. Here, we extend this approach to assess the WM benefits of refreshing different representation types, from colors to spatial locations and words. Across six experiments, we show that refreshing frequency modulates performance in all stimulus domains in accordance with the tenet that refreshing is a domain-general process in WM. The benefits of refreshing were, however, larger for visual-spatial than verbal materials.

DOI: https://doi.org/10.1111/nyas.13621

Originally published at:
Souza, Alessandra S; Vergauwe, Evie; Oberauer, Klaus (2018). Where to attend next: guiding refreshing of visual, spatial, and verbal representations in working memory. Annals of the New York Academy of Sciences, 1424(1):76-90.
DOI: https://doi.org/10.1111/nyas.13621

Research Article

Where to attend next? Guiding refreshing of visual, spatial, and verbal representations in working memory

Alessandra S. Souza [a], Evie Vergauwe [b], & Klaus Oberauer [a]

[a] University of Zurich, Switzerland

[b] University of Geneva, Switzerland

Short-title: Guiding Refreshing in Working Memory

Length = 8600 words (main text)

Keywords: attention; cuing; refreshing; working memory; short-term memory

Author Note

Alessandra S. Souza, and Klaus Oberauer, Department of Psychology, University of Zurich, Switzerland. Evie Vergauwe, Department of Psychology and Educational Sciences, University of Geneva, Switzerland. Correspondence should be addressed to Alessandra S. Souza, Department of Psychology, Cognitive Psychology Unit, University of Zürich, Binzmühlestrasse 14/22, 8050 Zurich, Switzerland. E-mail: a.souza@psychologie.uzh.ch

This article has been accepted for publication in Annals of the New York Academy of Sciences. It has undergone full peer review but has not been through the copyediting, typesetting, pagination, and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as: https://doi.org/10.1111/nyas.13621

Abstract

One of the functions that attention may serve in working memory (WM) is of boosting information accessibility, a mechanism known as *attentional refreshing*. Refreshing is assumed to be a domain-general process operating on visual, spatial, and verbal representations alike. So far, few studies have directly manipulated refreshing of individual WM representations to measure the WM benefits of refreshing. Recently, some of us developed a *guided-refreshing method*, which consists of presenting cues during the retention interval of a WM task to instruct people to refresh (i.e., attend to) the cued items. Using a continuous color reconstruction task, previous studies demonstrated that the error in reporting a color varies linearly with the frequency with which it was refreshed. Here, we extend this approach to assess the WM benefits of refreshing different representation types, from colors to spatial locations and words. Across 6 experiments, we show that refreshing frequency modulates performance in all stimulus domains in accordance with the tenet that refreshing is a domain-general process in WM. The benefits of refreshing were, however, larger for visual-spatial than verbal materials. The data and analysis scripts reported herein are available at the Open Science Framework (https://osf.io/skw8x/).

Introduction

Attention and working memory (WM) are interwoven cognitive processes with bidirectional relations: WM stores the representations that guide attention, [1–6] and attention helps managing the contents of WM. [7–10] The present article is concerned with the latter; in particular, with how attention shapes WM maintenance.

One of the functions that attention is assumed to serve for WM maintenance is increasing information accessibility, a process known as refreshing. Refreshing occurs when the focus of attention is directed to one representation in WM, leading to the strengthening of the binding between the representation and its retrieval cue, thereby facilitating subsequent access to this representation. [11–13] Refreshing is usually contrasted with articulatory rehearsal [14] which is assumed to be an articulation-based process for maintenance of speech-based (phonological) representations such as words, letters, or digits. One difference between articulatory rehearsal and refreshing is that rehearsal is blocked by concurrent articulatory activities (such as the constant repetition of irrelevant syllables), whereas refreshing is only prevented by attentionally demanding tasks [12,15].

Another difference between these two processes is that rehearsal applies only to verbal information, whereas refreshing is assumed to be domain-general. According to this tenet, attention can boost maintenance of different types of representations (visual, spatial, or verbal) in WM. So far, the evidence supporting this claim has come from studies assessing the costs of distracting attention during the retention interval of WM tasks. Requiring participants to perform an attentionally demanding concurrent task impairs memory for letters, [16–18] words, [17] spatial locations, [17,19,20] colors, [20–22] and color-shape bindings. [23] This impairment is a function of the relative amount of time spent on the distractor task over the total duration of the retention interval, a ratio known as the *cognitive load*. [24,25] Although the cognitive load effect is consistent with a domain-general refreshing process in WM, the evidence it provides is only indirect because distraction may impair WM for other reasons than by impeding refreshing. For example, processing of the stimuli in the distractor task creates representations of this information in WM, where it may interfere with the memoranda due to their overlap at the item level (e.g., shared phonological or semantic features) and at the context level (to which retrieval cue an item is associated). The cognitive-load effect could therefore also be explained by the use of attention to reduce the levels of interference due to removing of irrelevant information from WM. This is the explanation of the cognitive load effect implemented in a computational model of WM known as SOB-CS. [5] Distractor tasks may also impede other attentional processes apart from refreshing, e.g. by stopping the consolidation of an item in WM. [26–30]

To positively, and more directly, link refreshing to WM performance, some of us have developed an instructed-refreshing procedure to experimentally manipulate which items are refreshed in WM. [11,21] In these studies, participants memorized six colored dots. During the retention interval, cues (arrows) pointed to memory items, and participants were instructed to “think of” the cued item, hence bringing it to the focus of attention and refreshing it. The frequency (0, 1, or 2 times) with which items were cued for refreshing was varied. At the end of the trial, participants had to reproduce from memory the color of a test item, which was equally likely to have been refreshed 0, 1, or 2 times. Refreshing reduced the error in reproducing the color of the test item as a direct function of its refreshing frequency. This study provided first evidence that refreshing has a cumulative beneficial effect on visual WM.

This first set of studies showed that refreshing modulates the maintenance of colors in WM. This evidence, however, does not bear on the question of whether refreshing is domain general. Demonstrating that attending to different types of representations in WM improves their maintenance would provide corroborative evidence for this tenet. This is an important step for obtaining a better understanding of how refreshing works. So far, it is not clear how exactly refreshing operates on WM, and under which conditions it occurs. Knowing whether refreshing operates similarly across different representation domains constrains the space of hypotheses of how refreshing should be implemented in WM. Another reason for assessing whether a refreshing benefit can be obtained for different stimulus types relates to the claim that some types of representations may not be "refreshable". Some researchers have recently claimed that some visual and auditory materials – in particular unfamiliar, non-categorical ones – are non-refreshable. [17,31,32] The evidence to support this claim is that, for these materials, there is no trade-off between memory performance and concurrent task-processing. The instructed-refreshing manipulation provides an alternative way to directly assess whether or not a given stimulus material is refreshable. If some types of materials cannot be refreshed, guiding attention to these representations will not boost their maintenance.

Accordingly, the goal of the present set of studies was to use the instructed-refreshing manipulation to assess the WM benefits of refreshing across a broader range of stimulus types and task conditions (see Table 1). Our aim to use "think-of" cues to guide attentional refreshing to individual memory items hinges on the assumption that, in response to these cues, participants direct attention to the cued representations, and that doing so engages central attention. Evidence for these assumptions comes from a study combining a retro-cue with a tone discrimination task [33]. Retro-cue benefits were reduced when the stimulus onset asynchrony between the cue and the tone task was reduced. This is a classical psychological refractory period (PRP) effect, and it indicates that directing the focus of attention in working memory as prompted by a retro-cue requires (central) attention. This finding shows that retro-cues tap the same type of attention as investigated in the cognitive load effect.

To address the domain-generality of refreshing using the instructed refreshing paradigm, in the first two experiments (Exp. 1a and 1b), we assessed refreshing of spatial representations in conditions similar to the previous studies using colors (i.e., with simultaneously presented items). Next, we assessed whether refreshing could contribute to the maintenance of visual information (colors) even when this information is presented sequentially, as it is common in studies of verbal materials (Exp. 2). Finally, in the last series of experiments (Exp. 3a-c) we assessed refreshing of sequentially presented words.

Experiment 1

Experiment 1 was designed to test whether the refreshing effect generalizes to a dimension other than color. We chose the angular location of a dot in a ring [34,35] as the memoranda because they can be reported in a continuous scale akin to color. The experimental set-up was similar to the studies reported by Souza and colleagues [11,21]: The memoranda were presented simultaneously onscreen, followed by a brief retention interval wherein a sequence of cues was shown. Each cue instructed participants to refresh the cued item.

Testing memory for location posed a challenge in terms of how to cue the WM items. Previously, arrows were used to guide refreshing. Given that arrows contain angular information, they are likely to interfere with the memoranda. Hence we developed two alternative procedures to instruct refreshing. In Experiment 1a, the cue was a peripheral dot presented at the location of the WM item, whereas in Experiment 1b, the cue was a colored dot presented in the middle of the screen, with its color indicating the refreshing target (see

Figures 1a-b). The peripheral cue invites eye movements to the location of the indicated memory item, whereas central cueing discourages eye movements. It is unclear whether eye movements are desirable in the context of this task. On the one hand, some studies have shown that looking at the locations previously occupied by memory items is associated with improved memory retrieval. [36,37] On the other hand, eye movements can hamper spatial WM performance. [38] Given these two opposing possibilities, we ran these two task versions to explore the effectiveness of peripheral and central cues for guided refreshing.

Participants

Thirty students from the University of Zurich took part in Experiment 1a for three 1-hour sessions. In Experiment 1b, another set of thirty students took part in one session.

In all experiments reported here, participants read and signed an informed consent form prior to the study, and were debriefed at the end. Participants received financial reimbursement (15 Swiss francs per hour) or course credit. The experimental protocol is line with the ethical guidelines of the Institutional Review Board of the University of Zurich.

Materials and Methods

Participants performed a continuous location reconstruction task [26,35] under three conditions: Baseline, Refreshing, and Dual-task. We will first describe the general task as implemented in the Baseline condition, followed by the modifications of this baseline for the implementation of the refreshing manipulation, and also of a dual-task situation (Experiment 1a, data not reported here). In both experiments, the memoranda comprised the angular location (1° to 360°) of a dot presented at the edge of a ring (see Figures 1a and 1b). There were five colored rings in each memory array (evenly spaced on an imaginary circle). Color (yellow, pink, green, red, and blue) indicated the spatial position of the ring on the screen, and was invariant. We used five colors to assist discrimination of the different items and to minimize grouping of the dots into a single figure. Each trial started with the presentation of the colored rings without dots for 1 s, followed by the simultaneous presentation of the dots in each ring for another 1 s. Afterwards the retention interval commenced (duration = 3 s). In Experiment 1a, black rings serving as placeholders were presented throughout the retention interval. In Experiment 1b, all rings were removed from the screen during the retention interval to reduce any incentive to glance back at the locations previously occupied by the memoranda. In the Baseline condition, the retention interval was not filled by additional processing demands. At the end of the retention interval, participants were prompted to reproduce the orientation of a target item (indicated by the unique colored ring) by clicking on a point on the colored ring. Only responses within the ring edge were accepted. A blank 2-s inter-trial interval followed each response. This Baseline condition provides an overall assessment of recall performance in a situation in which participants are free to maintain the memoranda as they wish, and in which interference during the retention interval is minimal.

The Refreshing condition was identical to the Baseline except that a sequence of four cues was presented during the retention interval. The cue was a peripheral black dot presented in the center of one placeholder ring in Experiment 1a, and a colored dot in the screen center in Experiment 1b (see Figures 1a-b). Participants were instructed to "think of" the WM item indicated by the cue (hereafter referred to as a refreshing step). Participants were also instructed that the cues did not reliably indicate the test item. Two sucessive cues always indicated different items. The cue sequence could point to 4 different items, or sometimes one or two items would be cued twice, resulting in three or only two items being refreshed in a trial. The procedure for generating cue sequences was the same as implemented by Souza and colleagues [11,21] and is described in detail in the Online

Supplementary Materials. Overall, each item in the memory array was cued either 0, 1, or 2 times during the retention interval. We aimed to select the target of recall in an equal proportion of trials to be a 0-, 1-, or 2-Refreshing item (with these conditions being randomly intermixed). In Experiment 1a, due to a programming error, targets were selected with an equal probability for all six items, and hence 2-Refreshed items (1 out of 5 items) were tested less often than 0- and 1-Refreshing items (which had similar testing probability).

In Experiment 1a participants additionally completed sessions in which a secondary task was carried out during the retention interval. The secondary task comprised a tone-pitch discrimination (high or low; tone duration = 75 ms). Participants had to press the upper or the lower arrow key for a high or low tone, respectively. The number of tones (1, 2, or 3) presented during the retention interval was varied randomly across trials to assess the impact of imposing different levels of cognitive load. [25] Tones were evenly spaced during the retention interval, with the first tone always occuring 0.1 s after the offset of the memory array. Here we focus mainly on the results of the Refreshing and Baseline conditions. The interested reader can find the detailed results of the cognitive-load manipulation in the Online Supplementary Materials (and the raw data at the OSF). [a]

We blocked all conditions to facilitate instruction and to allow participants to prepare for the tasks to the best of their abilities. In Experiment 1a, participants completed 100 trials

[a] Our original intention was to investigate the relation between the cognitive load effect and the refreshing frequency effect across different stimulus domains. However, we did not find evidence for a cognitive-load effect on the recall of spatial locations and of words (Experiment 3). Given that the cognitive load has been extensively replicated with verbal memoranda, it is unclear whether the lack of the cognitive load observed here is related to characteristics of the recall task, of the distractor task, or their combination. Future experiments are required to examine why this is the case. Hence we have refrained from speculating on this boundary condition of the cognitive-load effect here, and we deferred the examination of the relation of the cognitive-load effect and refreshing-frequency effect to future studies.

of the Baseline, 300 trials of the Refreshing condition, and 300 trials of the Dual-task condition. The Refreshing condition was completed in one separate session, whereas the Baseline and Dual-task conditions were completed in different blocks in the remaining two sessions. In Experiment 1b, participants completed a single session comprising one block of 75 trials of Baseline and one block with 225 trials of the Refreshing condition. The order of the blocks within a session, and session order, was fully counterbalanced across participants.

Data analysis

We computed the absolute distance between the reported angle and the true angle of the target (ranging from 0 to 180°; hereafter *recall error*). We submitted this data to a Bayesian analysis of variance (BANOVA) [39], using the BayesFactor package [40] implemented in R. The BANOVA implements general linear models that include the main effects of the predictors and their interactions in several combinations. The likelihood of each of these alternative models (e.g., M_1 = main effects of A and B) is compared against a null model (M_0). The ratio of their likelihoods is the Bayes Factor (BF). The BF provides a factor by which our ratio of prior beliefs in the two models should be updated in light of the data. To assess the evidence for each predictor in a model, one computes the ratio of the BF of the model including the predictor vs. the model omitting it, leaving everything else equal. For instance, to assess the evidence for a two-way interaction, a model including both main effects and their interaction is compared to a model with the two main effects but without the interaction. Here we report the model with the highest likelihood against the Null, and we assess the evidence for each individual predictor included in this model by computing the BF for keeping vs. removing that predictor from the winning model.

We also submitted the data of the Refreshing condition to mixture modeling using the CatContModel package. [41] We fitted the Zhang and Luck model [42], which assumes that responses come from a mixture of distributions representing information retrieved from memory (P^M) and guessing ($1 - P^M$). The model futher assumes that the variability of the responses around the target value (σ) reflects memory imprecision. We fitted the model (30000 iterations; 5000 burn-in) separately for each experiment and then combined their posterior distributions to assess their combined evidence for an effect of refreshing frequency, and for an effect of cue-type. We compared models which included an effect of refreshing frequency on both P^M and σ (M_1), or in only one of these parameters (M_2 and M_3, respectively), using the Watanabe-Akaike Information Critetion (WAIC). We selected the best model (lower WAIC; see Table 2), and then we evaluated the BF for the effects of the predictors on the parameter estimates included in this model.

Results

First, we pre-screened the data for outliers. Outliers were defined as participants showing an average recall error two absolute standard deviations (MAD) above the group median. [43] There were no outliers in Exp.1a (group median = 39.4°, MAD = 12.2°). Two participants in Exp. 1b were identified as outliers and were excluded from subsequent analyses: one had a median recall error of 94.5° and the other of 75° (group median = 45°, MAD = 13.8°). Note that average performance around 90° is indicative of pure guessing.

Refreshing Frequency Effect. The effect of refreshing frequency on recall error (see Figure 2a) was statistically supported over the Null in each experiment, Exp.1a, BF_{10} = 8.5; Exp. 1b, BF_{10} = 23.6. A combined analysis across experiments having refreshing frequency and cue type (peripheral vs. central) as predictors showed that the best model included the main effects of refreshing frequency and cue type (BF_{10} = 3592.87). The evidence for the refreshing effect was overwhelming in this analysis (BF_{10} = 2050.1), whereas the effect of cue type was ambiguous (BF_{10} = 1.38). There was evidence against including their interaction in the model (BF_{10} = 0.21), suggesting that the two cue types were equally effective in guiding refreshing. We re-ran the BANOVA restricting the number of refreshing levels to two to compare adjacent refreshing levels: there was some evidence of improvement from 0 to 1 refreshing steps (BF_{10} = 2.8), and strong evidence for an improvement from 1 to 2 refreshing steps (BF_{10} = 17.81).

We compared several mixture models in which different parameters were allowed to vary with refreshing frequency (see M_1 to M_3 in Table 2). The best model definitely included the effect of refreshing frequency on probability of recall (see Figure 2b). Whether the effect of refreshing on memory imprecision should be included in the model was more ambiguous. For Exp. 1a, including an effect in this parameter did not improve the model, whereas it did for Exp. 1b. The combined analysis across both experiments indicated that the best model included only an effect of refreshing frequency on probability of recall (M_2). Hence, we selected this model for further analysis. The evidence for a refreshing frequency effect on recall probability was substantial in Exp. 1a (BF_{10} = 7.4) and strong in Exp. 1b (BF_{10} = 20.5). For their combined analysis, the evidence was very strong (BF_{10} = 100.2) for a refreshing frequency effect on recall probability. The evidence was somewhat against an effect of cue type (BF_{10} = 0.6), and strongly against its interaction with refreshing frequency (BF_{10} = 0.006).

Instructed Refreshing vs. Baseline. In the Baseline condition, the retention interval was blank and participants could, in principle, engage in spontaneous refreshing. Furthermore, in the Baseline, interference is minimal, hence allowing us to gauge whether presenting the refreshing cues is disruptive to memory. We compared performance in the Baseline condition against the 0- and 2-Refreshing conditions: there was overwhelming evidence for worse recall in the 0-Refreshing condition (BF_{10} = 3189.8), and also some evidence for worse recall in the 2-Refreshing condition (BF_{10} = 2.63). In both analyses, there was no evidence for an interaction with cue type (BF_{10} = 0.33; BF_{10} = 0.27), and the evidence for a main effect of cue type was ambiguous (BF_{10} = 1.54; BF_{10} = 0.95). This result suggests that although refreshing modulated recall, as reflected by the refreshing frequency effect, participants could perform better in the task when they were left on their own during the retention interval. This result could indicate that presentation of four cues during the retention interval disrupts memory for spatial locations. Alternatively, this result could mean that, for spatial-location memoranda, refreshing is not actually beneficial, but only impairs access to the subset of items not refreshed.

Discussion

Experiment 1 showed a refreshing frequency effect with continuous spatial locations, extending this effect to another stimulus dimension than color, and indicating that these representations can receive an attentional boost. Mixture modeling revealed that refreshing affected recall probability, but not precision, replicating previous reports with color stimuli.[11,21] The refreshing frequency effect was independent of cuing mode (peripheral or central). The peripheral cue, which could have prompted more eye movements, did not hinder spatial WM. If anything, performance in this task tended to be better than in the central color-cue version. In our study, however, we have not tracked eye-movements and hence we can only speculate on whether the central cueing condition indeed reduced the incentive for glancing back at the locations of the items. Although processing of the cue itself did not require eye-movements, it is still conceivable that participants may have spontaneously glanced back at the locations of the cued items. Future studies should therefore address this question by recording eye movements directly, and instructing participants not to move their eyes.

Overall performance in the Refreshing condition was worse than in the unfilled Baseline. By contrast, in the studies with color stimuli reported by Souza and colleagues [11,21], recall of 2- Refreshing targets tended to be similar or better than recall in the unfilled Baseline. The observation of worse recall for 2-Refreshing targets for spatial information is intriguing and may indicate that the refreshing instruction introduced some sort of unwanted interference. A reviewer raised the possibility that the refreshing cues could have prematurely stopped encoding or consolidation into visual WM. Previous work has shown that encoding into visual WM takes about 50 ms per item [29]. Given that we presented the memory array for 1000 ms, and the first refreshing cue only appeared after an additional interval of 500 ms, it is unlikely that the refreshing cues impaired memory by disrupting this process.

It remains, however, possible that eye-movements induced by the refreshing cues could explain this cost. As we outlined above, our effort to use central cues may not have prevented participants from moving their eyes. Yet another possible explanation for this finding is that participants engage in other forms of maintenance than item-level refreshing when left on their own. For example, participants may engage in elaborative rehearsal or visual imagery of a single figure comprising all dots in the entire array, which in turn could be more advantageous for the maintenance of this information. Testing the viability of these conjectures was beyond the scope of this research project.

Experiment 2

The goal of Experiment 2 was to examine whether the refreshing frequency effect is also observed when the visual memoranda are presented sequentially for encoding. When items are presented sequentially, the order of the refreshing cues will hardly ever match the order of presentation of the items. Some authors have proposed that refreshing naturally proceeds in forward serial order. [12,17,44–47 but see 48,49] If that is the case, the refreshing cues would conflict with this natural tendency. To the degree that participants can still use spatial locations to reactivate the cued representation, a refreshing frequency effect should be observed. This is an important first step in bridging between modes of testing of visual and verbal information, because for verbal information sequential encoding of items is inevitable (e.g., reading of words).

Materials and Method

A new sample of participants (n = 24) completed two sessions of a continuous color reproduction task. [36,44,45] Each trial began with a message requiring participants to start with the continuous repetition of *"bababa"* throughout the trial (articulatory suppression, AS), and to press the spacebar to initiate the trial. AS prevents verbal labeling of the colors, which can aid their maintenance in WM. [52] We implemented AS in Experiment 2 because this experiment was planned for a direct comparison with Experiment 3 using verbal stimuli, and therefore we wanted to make sure that participants in Experiment 2 used visual representations to maintain the stimuli. In our previous studies using color stimuli we have observed a refreshing-frequency effect both with AS [11] and without AS [21], and therefore we expect to find a refreshing-frequency effect in the present experiment, similar to Experiment 1.

After pushing the key to start the trial, a white fixation cross appeared for 0.5 s. Then, 5 colored dots were presented sequentially in different locations on a virtual circle, each shown for 0.4 s (see Figure 1c). Colors were selected from 360 values distributed along a color circle (radius = 60) defined in the CIELAB color space, with L = 70, a = 20, and b = 38. [42] The colors of the memoranda were selected to be at least 20° apart from each other on the color circle. The retention interval had a total duration of 3 s. The sequence of refreshing cues started 0.5 s after the offset of the memoranda, and finished 0.5 s before the onset of the memory test. Each cue was shown for 0.5 s, and consisted of a central white arrow pointing to the location of a memory item. Cue sequences were as in Experiment 1. At the end of the retention interval, a white circle indicated the location of the test item, and participants were instructed to reconstruct the color of the test item by clicking on a point in a color wheel. Participants completed a total of 480 trials, which were evenly distributed across two sessions, and the three refreshing frequency levels. This experiment did not include a Baseline condition without a refreshing manipulation, because it was not necessary for testing whether the refreshing-frequency benefit can be obtained with sequential presentation.

Results

One participant was identified as an outlier with a median recall error of 86.2° (group median = 55.5, MAD = 14.9), and was therefore excluded from subsequent analyses.

Refreshing Frequency Effect. There was strong evidence (BF_{10} = 22.35) for an effect of refreshing frequency on recall error (see Figure 3a). Comparison of individual refreshing levels with other each using t-tests yielded substantial evidence for a reduction in recall error from 0 to 1 refreshing steps (BF_{10} = 6.41), but the evidence was ambiguous for a further reduction from 1 to 2 refreshing steps (BF_{10} = 0.99). Mixture modeling showed that the best model included an effect of refreshing only for the probability of recall, but not for memory imprecision (see Table 2 and Figure 3b), and the BF for the main effect of refreshing on recall probability was 6.2.

Discussion

Experiment 2 showed that refreshing instructions directed at individual items in random order can be applied to visual items even when these are presented sequentially. In the case of sequential presentation of the memoranda, many theorists would assume that refreshing proceeds serially and in forward order, reproducing the serial input position of the items. [12,17,44–46] Hence, the refreshing instruction would be in direct conflict with the "natural" order of spontaneous refreshing. This sort of conflict could have reduced the probability that participants followed the refreshing instruction, thereby diluting the refreshing frequency effect. In contrast to this possibility, we observed a healthy refreshing frequency effect, replicating and extending our previous work with the simultaneous presentation of colored dots, and also the results of Exp. 1. Hence, our study shows that randomly ordered cues can also be used to guide refreshing in conditions in which the information is presented serially. This builds an important bridge for testing of the refreshing frequency effect with verbal materials, for which sequential presentation is a standard procedure.

Experiment 3

In the final series of experiments, we tested whether guided refreshing yields better recall of words from verbal WM. Tests of verbal WM usually comprise the serial presentation of the memoranda for an immediate forward serial recall test. This procedure differs in many regards from tests of visual WM. As discussed in the context of Experiment 2, one difference pertains to the sequential presentation of the items. Experiment 2 showed that this difference is, however, not critical to obtaining a refreshing frequency effect. Another difference refers to the requirement to recall all items in forward serial order. As we already pointed out, the refreshing frequency effect was obtained with cue sequences in which order of item cueing was random. This is probably not ideal for the maintenance of order information. Hence, as a first step towards establishing a refreshing frequency effect with verbal materials, we bridged between the procedures for testing visual WM and verbal WM by using a single probed recall test akin to the one used in tests of visual WM.

We conducted three experiments that differed regarding two main aspects. First, in Experiments 3a and 3b participants performed AS (constant articulation of *"bababa"*), thereby preventing the use of articulatory rehearsal for the maintenance of the words; this requirement was removed in Exp. 3c. It has been argued that articulatory rehearsal and refreshing can be used additively to maintain verbal information in WM [15,47,53], and that blocking of rehearsal would force participants to rely more strongly on refreshing [17]. Hence we explored whether AS would change the boost provided by the refreshing instruction. Second, in Experiment 3a recall was oral whereas in Experiments 3b and 3c it was typed (see Table 1). These variations were intended to explore the boundaries of the refreshing frequency effect with verbal memoranda.

Participants

Thirty-nine students from the University of Zurich took part in Experiment 3a for one session. Two new samples of students took part in Experiments 3b (n = 24) and 3c (n = 24) for two sessions.

Materials and Method

In the beginning of a trial, six white boxes arranged in a circle were shown against a black (Exp. 3a) or grey background (Exps. 3b and 3c) for 1 s. Next, one at a time, a word appeared in one of the boxes (for 1 s), starting with the top-left box and proceeding in clockwise order (see Figure 1D). Words were selected from a pool of 622 German nouns with 4-5 letters. Words were sampled from this pool without replacement until all words were used once. Then, words were sampled anew without replacement. After the last word, the retention interval started (2.5 s in Exp. 3a, and 3 s in Exps. 3b and 3c). In the Baseline condition, only the box placeholders were visible during the retention interval. In the Refreshing condition, white central arrows were presented, each pointing in the direction of one memory item for 0.5 s. Cue sequences and refreshing instructions were the same as in Experiment 1. In Experiments 3b and 3c, the cue sequence was preceded by a 0.5 pre-cue interval and followed by a 0.5 s post-cue interval, as in Experiments 1 and 2. At the end of retention interval, a single probed recall test followed: one of the placeholders turned yellow, and a question mark appeared therein (indicating the target item), whereas all remaining placeholders turned light grey. In Experiment 3a, participants were instructed to say aloud the word that was presented at the probed location. In Experiments 3b and 3c, the recall mode was changed to typed. The typed word appeared in the center of the screen in yellow. Participants could correct their input using the backspace, and then they pressed the Enter key to confirm their response. A blank screen was inserted for 2 s before the next trial started. The three experiments differed regarding the instructions to perform AS concurrently: in Exps. 3a and 3b participants repeated the syllabus *"bababa"* throughout the trial; this requirement was removed in Exp. 3c. In addition to Baseline and Refreshing trials, in Experiments 3b and 3c, participants completed a Dual-task condition akin to the one in Experiment 1a (tone discrimination task with 1, 2, or 3 tones). We note that we did not find evidence for an effect of cognitive load on recall of words in any of these experiments; detailed results can be found in the Supplementary Materials.

In all experiments participants completed 36 trials per design cell. In Exp. 3a, Baseline and Refreshing trials (with 0-, 1-, and 2-Refreshing targets) were randomly intermixed. In Exp. 3b and 3c, Baseline and Refreshing trials were completed in separate blocks within the same session. Dual-task trials were completed in a different session. Session order (and condition order within a session) was counter-balanced across participants.

Results and Discussion

The dependent measure in this task was proportion of correctly recalled words. In the typed recall version, the data was automatically scored by the program as correct if there was a perfect match between the target word and the typed response. To consider possible typos, we first computed the distance between the response and the target word using the Levenshtein metric. The Levenshtein distance is a metric of how many changes (additions or deletions) would be necessary to transform one word-string (e.g., the response) to the other string (target). We selected responses with a distance of 1 and 2, and checked whether they reflected typos. We considered as typos any addition or deletion of characters which would not change the meaning of the intended target word. We corrected the accuracy score when it was clear that participants remembered the correct word but made a spelling mistake; e.g., “Norden” when presented with “Nord”; e.g., “Kurvee” instead of “Kurve”, or when they entered a plural version, e.g., “Augen” instead of “Auge”.

Next, the data was prescreened for outliers (performance 2*MAD below the group median). No outliers were identified in any of the experiments: Exp. 3a, group median = 0.43 and MAD = 0.15, Exp. 3b group median = 0.51 and MAD = 0.15, and Exp. 3c group median = 0.69 and MAD = 0.21.

Refreshing frequency effect. Overall, accuracy tended to vary linearly with refreshing frequency (see Figure 4). We computed the evidence for an effect of refreshing frequency in each individual experiment: Exp. 3a (BF_{10} = 1.5), Exp. 3b (BF_{10} = 0.44), and Exp. 3c (BF_{10} = 7.85). This analysis suggests that the refreshing frequency effect was ambiguous when participants had to perform AS concurrently (Exps. 3a and 3b). Next, we computed the evidence for the refreshing frequency effect across all experiments entering refreshing frequency, AS, and recall mode as categorical predictors in a BANOVA. The best model included only the main effects of refreshing frequency and AS (BF_{10} = 7.9 x 10^8). Including recall mode did not improve the model (BF_{10} = 0.96). This indicates that the somewhat larger performance for typed in comparison to oral recall is statistically not credible. There was strong evidence to keep both the refreshing frequency effect (BF_{10} = 120.7) and the AS effect (BF_{10} = 10069) in the winning model. There was however evidence against including an interaction between refreshing frequency and AS (BF_{10} = 0.11), and of refreshing frequency and recall mode (BF_{10} = 0.07). Overall, these results suggest that with a sufficiently large sample size (across the three experiments, n = 87), we can obtain strong enough evidence for an effect of refreshing with verbal materials. With this large sample size, we also compared adjacent refreshing levels with each other by re-running the BANOVA models with only two refreshing levels. There was evidence supporting an improvement from 0 to 1 (BF_{10} = 5.5), but not from 1 to 2 (BF_{10} = 0.51).

Instructed Refreshing vs. Baseline. First, we assessed whether overall performance (across all three experiments) differed between the Baseline and Refreshing condition, but the evidence was against an effect of condition (BF_{10} = 0.26). Second, we compared 0-Refreshing items to recall in the Baseline. There was evidence against worse performance for 0-Refreshed items (BF_{10} = 0.33). Third, we compared 2-Refreshing items to the Baseline: There was tentative evidence for better recall of 2-Refreshing items (BF_{10} = 2.21).

Discussion

When taking the data of the three experiments together, there was strong evidence for a refreshing frequency effect, even though this evidence was mostly weak when considering each experiment separately. This was particular the case when the experiments included AS, which is assumed to block articulatory rehearsal. Some authors have argued that articulatory rehearsal and refreshing can be used additively to maintain verbal information in WM [15,47,53], and that blocking of rehearsal would force participants to rely more strongly on refreshing. [17] Here, if anything, the refreshing frequency effect was more substantial in the absence of AS (Exp. 3c). In the combined analysis across the three experiments, there was however no evidence for an interaction between AS and refreshing frequency. This indicates that AS did not change the use of the refreshing instruction, but the strong memory impairment induced by it reduced the levels of performance so drastically that the effect of refreshing was more difficult to detect.

For verbal materials, performance in the Baseline was not better than in the Refreshing condition. If anything, there was a tendency for better performance in the 2-Refreshing condition in comparison to the Baseline. This is similar to the general pattern obtained with colors [11,21], but sets it clearly apart from the continuous location information (Exps. 1a and 1b). These results indicate that for verbal information, the refreshing instruction was not disruptive (it yielded minimal interference), and that we may even push performance to levels higher than those observed when participants are left to their own.

Refreshing Effect Across All Experiments

Finally, we examined the size of the refreshing frequency effect across all experiments reported here. To be able to do so, we took as dependent variable the probability of correctly retrieving the target item as estimated from the mixture model applied to continuous-reproduction responses, which is analogous to the proportion correct measure computed for words. Figure 5 presents the posterior of the increase in performance obtained from 0 to 2-Refreshing steps, separately for the visual-spatial (locations and colors) and verbal (words) memoranda. Figure 5 shows similar refreshing boosts for spatial locations and colors, which were larger than the ones obtained for words.

However, the size of the refreshing effect for visual and verbal materials did overlap to some extent, particularly when no AS was imposed on the concurrent maintenance of the words.

General Discussion

Refreshing is usually conceived as a domain-general process operating on visual, spatial, and verbal representations in WM alike. [12,13,31,54] So far, the putative role of refreshing across different stimulus domains has only been examined by imposing concurrent task demands, and assessing the costs incurred (i.e., the cognitive load effect). Here, we looked at the positive aftereffects of attending to information in WM. We used cues to instruct participants to refresh continuous spatial locations, continuous colors, and words. Across all experiments, we were able to demonstrate a monotonic improvement in the recall of these materials from WM as a direct function of their refreshing frequency.

Refreshing as a Cumulative Focusing Benefit

We see refreshing as an increase in the accessibility of a representation after it enters the focus of attention, which remains after the focus moves away. [55] If this was not the case, refreshing could only benefit one item at a time, instead of yielding a cumulative benefit. Hence, the refreshing effect we measured reflects the final product of several refreshing steps distributed over the course of the retention interval, with each refreshing step conveying a persistent boost to the accessibility of the refreshed representation. Here we provided a first assessment of the impact of using cues to guide the sequential and gradual strengthening of different types of representations in WM, and we provide a first comparison of the relative size of this boost for different stimulus materials.

Domain-General Refreshing

Our findings are in line with a cumulative refreshing effect for spatial locations, colors, and words. This is in line with the assumption that refreshing is domain-general. The reasoning for assessing a refreshing frequency effect across different stimulus materials is simple: If a representation is boosted by focusing attention on it, then this representation is refreshable. In our view, this provides the strongest test of whether a representation can or cannot be refreshed. In contrast, previous studies have relied on the absence of a cognitive-load effect, or on the observation of forgetting over the course of an unfilled interval, to claim that representations are not refreshable. [17,31] Those findings can, at best, speak to the hypothesis that participants spontaneously refresh the memoranda during periods of free time, and not to the refreshable nature of a representation. Conversely, the instructed refreshing procedure can only demonstrate whether a kind of representation is refreshable, not whether people spontaneously refresh them. Future studies investigating whether refreshing occurs for certain types of memoranda should take into consideration these different methodologies and the types of questions they can address.

Does Refreshing Operate Alike for All Types of Memoranda?

The refreshing frequency effect was larger for visual-spatial materials than for words. Why was that the case? We don’t have a ready answer to this question. There are several possibilities to consider. One possibility is that the refreshing instruction is less effective for verbal materials because verbal lists are presented sequentially, and participants may therefore naturally attempt to refresh these items in cumulative forward serial order. This is unlikely to be an explanation for several reasons. First, Experiment 2 showed a sizeable refreshing frequency effect with visual materials presented sequentially. Second, a recent study [56] compared the effectiveness of cumulative forward-order refreshing of verbal lists versus other refreshing schedules that differed systematically from it. There was no difference in performance across all instructed refreshing orders, indicating that cumulative forward-order refreshing is not better for the maintenance of verbal lists. Third, simulations using a computational model of WM has shown that data from complex span tasks and the cognitive load effect can be better reproduced when flexible refreshing schedules are implemented. [48,49]

Another possibility relates to strategic use of the cues. At the moment, we have no means of determining whether participants were indeed following the refreshing instruction other than by observing its final effect on memory accuracy. Hence, it could be that participants were more reluctant to follow the refreshing instruction with verbal than visual-spatial materials. This may be the case because participants are more used to apply strategies to the retention of verbal information in their daily tasks (e.g., to remember a shopping list, a message for a friend, a PIN number, etc) which may conflict with the experimentally imposed instruction. This highlights the importance of further developing ways of determining whether people are following the refreshing cues to the best of their abilities. Recently, it has been proposed that the refreshing-instruction manipulation may be useful to investigate the more local effect of refreshing on the speed with which people can accept a memory probe (a letter) as being part of the memory list. In this task, the last-refreshed letter was always recognized faster than any other letter in the list, providing independent evidence that participants followed the refreshing instruction. [46] It may be possible to combine both an assessment of the immediate recognition speed and of the cumulative effects of refreshing on recall accuracy in the same study to examine whether the two indicators of refreshing – faster recognition and improved recall accuracy – converge on a weaker effect of refreshing instructions for verbal than for visual materials.

Refreshing Effect on Mixture Model Parameters

Our experiments with continuous visual-spatial stimuli converged on refreshing having an impact on the probability of recalling the target, but weak or no evidence for an effect on memory precision. This observation is consistent with our two previously published studies with the refreshing instruction [11,21]. This pattern of findings is also in line with the current state of the literature on the effects of single retro-cues on continuous visual WM tasks (for a review see [10]), and with the effect of cognitive load found in this task [57]. We note, however, that there is an ongoing debate regarding how to best model responses in continuous reproduction tasks. [57–59] The mixture model used [42] here, although one of the most popular, may not be the best descriptor of all of the cognitive processes involved in performance in this task. [60] Hence we caution against over-interpreting the psychological meaning of the estimated parameters.

Conclusion

Using a guided-refreshing method, we demonstrated that refreshing frequency modulates memory accuracy for different stimulus types in WM, from colors to spatial locations and words. The observed boost was, however, larger in the visual-spatial than verbal domain. The guided-refreshing method provides one complementary tool to understand the positive aftereffects of attending to information in WM, and of assessing how and under which conditions refreshing is beneficial.

Acknowledgment

This research was supported by a grant from the Swiss National Science Foundation (project 149193) to K. Oberauer. All authors contributed to the study design. A. Souza programmed the experiments, analyzed the data, and wrote the first draft of the manuscript. The remaining authors provided critical feedback and edited the manuscript accordingly.

Conflict of Interest

The authors declare no conflicts of interest in publishing this research.

References

1. Desimone R. & J. Duncan. 1995. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**: 193–222.
2. Olivers C.N. L. 2008. Interactions between visual working memory and visual attention. *Front. Biosci.* **13**: 1182.
3. Olivers C.N.L., J. Peters, R. Houtkamp, *et al.* 2011. Different states in visual working memory: when it guides attention and when it does not. *Trends Cogn. Sci.* **15**: 327–334.
4. Kiyonaga A. & T. Egner. 2012. Working memory as internal attention: Toward an integrative account of internal and external selection processes. *Psychon. Bull. Rev.* **20**: 228–242.
5. Oberauer K., S. Lewandowsky, S. Farrell, *et al.* 2012. Modeling working memory: an interference model of complex span. *Psychon. Bull. Rev.* **19**: 779–819.
6. Kiyonaga A. & T. Egner. 2014. The Working Memory Stroop Effect: When Internal Representations Clash With External Stimuli. *Psychol. Sci.* **25**: 1619–1629.
7. Gazzaley A. 2011. Influence of early attentional modulation on working memory. *Neuropsychologia* **49**: 1410–1424.
8. Gazzaley A. & A.C. Nobre. 2012. Top-down modulation: bridging selective attention and working memory. *Trends Cogn. Sci.* **16**: 129–135.
9. Myers N.E., M.G. Stokes & A.C. Nobre. 2017. Prioritizing Information during Working Memory: Beyond Sustained Internal Attention. *Trends Cogn. Sci.* **21**: 449–461.
10. Souza A.S. & K. Oberauer. 2016. In search of the focus of attention in working memory: 13 years of the retro-cue effect. *Atten. Percept. Psychophys.* **78**: 1839–1860.
11. Souza A.S., L. Rerko & K. Oberauer. 2015. Refreshing memory traces: thinking of an item improves retrieval from visual working memory. *Ann. N. Y. Acad. Sci.* **1339**: 20–31.
12. Barrouillet P. & V. Camos. 2012. As time goes by temporal constraints in working memory. *Curr. Dir. Psychol. Sci.* **21**: 413–419.
13. Johnson M.K. 1992. MEM: Mechanisms of recollection. *J. Cogn. Neurosci.* **4**: 268–280.
14. Baddeley A. 1986. Working memory. *Clarendon Press. Univ. Press*.
15. Camos V. 2015. Storing Verbal Information in Working Memory. *Curr. Dir. Psychol. Sci.* **24**: 440–445.
16. Vergauwe E. & N. Cowan. 2014. Attending to items in working memory: evidence that refreshing and memory search are closely related. *Psychon. Bull. Rev.* **22**: 1001–1006.
17. Vergauwe E., V. Camos & P. Barrouillet. 2014. The impact of storage on processing: How is information maintained in working memory? *J. Exp. Psychol. Learn. Mem. Cogn.* Advance online publication.
18. Barrouillet P., S. Bernardin, S. Portrat, *et al.* 2007. Time and cognitive load in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **33**: 570–585.
19. Vergauwe E., P. Barrouillet & V. Camos. 2010. Do mental processes share a domain-general resource? *Psychol. Sci.* **21**: 384–390.
20. Vergauwe E., P. Barrouillet & V. Camos. 2009. Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *J. Exp. Psychol. Learn. Mem. Cogn.* **35**: 1012–1028.
21. Souza A.S. & K. Oberauer. 2017. The contributions of visual and central attention to visual working memory. *Atten. Percept. Psychophys.* **79**: 1897–1916.
22. Kiyonaga A. & T. Egner. 2014. Resource-sharing between internal maintenance and external selection modulates attentional capture by working memory content.,

Resource-sharing between internal maintenance and external selection modulates attentional capture by working memory content. *Front. Hum. Neurosci. Front. Hum. Neurosci.* **8**, **8**: 670–670.
23. Vergauwe E., N. Langerock & P. Barrouillet. 2014. Maintaining information in visual working memory: Memory for bindings and memory for features are equally disrupted by increased attentional demands. *Can. J. Exp. Psychol. Can. Psychol. Expérimentale* **68**: 158–162.
24. Barrouillet P., S. Bernardin & V. Camos. 2004. Time constraints and resource sharing in adults' working memory spans. *J. Exp. Psychol. Gen.* **133**: 83–100.
25. Barrouillet P., S. Portrat & V. Camos. 2011. On the law relating processing to storage in working memory. *Psychol. Rev.* **118**: 175–192.
26. Ricker T.J. & K.O. Hardman. submitted. The nature of short-term consolidation in visual working memory. *Manuscr. Submitt. Publ.*
27. Ricker T.J. 2015. The role of short-term consolidation in memory persistence. *AIMS Neurosci.* **2**: 259–279.
28. Bayliss D.M., J. Bogdanovs & C. Jarrold. 2015. Consolidating working memory: Distinguishing the effects of consolidation, rehearsal and attentional refreshing in a working memory span task. *J. Mem. Lang.* **81**: 34–50.
29. Vogel E.K., G.F. Woodman & S.J. Luck. 2006. The time course of consolidation in visual working memory. *J. Exp. Psychol. Hum. Percept. Perform.* **32**: 1436–1451.
30. Xu Y. 2017. Reevaluating the Sensory Account of Visual Working Memory Storage. *Trends Cogn. Sci.* **21**: 794–815.
31. Ricker T.J. & N. Cowan. 2010. Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *J. Exp. Psychol. Learn. Mem. Cogn.* **36**: 1355–1368.
32. Nees M.A., E. Corrini, P. Leong, *et al.* 2017. Maintenance of memory for melodies: Articulation or attentional refreshing? *Psychon. Bull. Rev.* 1–7.
33. Janczyk M. & M.E. Berryhill. 2014. Orienting attention in visual working memory requires central capacity: Decreased retro-cue effects under dual-task conditions. *Atten. Percept. Psychophys.* **76**: 715–724.
34. Ricker T.J. & K.O. Hardman. 2017. The nature of short-term consolidation in visual working memory. *J. Exp. Psychol. Gen.* **146**: 1551–1573.
35. Rouder J.N., J.E. Thiele & N. Cowan. 2014. Evidence for guessing in working-memory judgments. *Pap. Present. 55° Annu. Meet. Psychon. Soc. Long Beach Calif.*
36. Ferreira F., J. Apel & J.M. Henderson. 2008. Taking a new look at looking at nothing. *Trends Cogn. Sci.* **12**: 405–410.
37. Scholz A., A. Klichowicz & J.F. Krems. 2017. Covert shifts of attention can account for the functional role of "eye movements to nothing." *Mem. Cognit.* 1–14.
38. Postle B.R., C. Idzikowski, S.D. Sala, *et al.* 2006. The selective disruption of spatial working memory by eye movements. *Q. J. Exp. Psychol.* **59**: 100–120.
39. Rouder J.N., R.D. Morey, P.L. Speckman, *et al.* 2012. Default Bayes factors for ANOVA designs. *J. Math. Psychol.* **56**: 356–374.
40. Morey R.D. & J.N. Rouder. 2014. "*BayesFactor: Computation of Bayes factors for common designs*."
41. Hardman K.O. 2016. "*CatContModel: Categorical and Continuous Working Memory Models for Delayed Estimation Tasks*."

42. Zhang W. & S.J. Luck. 2008. Discrete fixed-resolution representations in visual working memory. *Nature* **453**: 233–235.
43. Leys C., C. Ley, O. Klein, *et al.* 2013. Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *J. Exp. Soc. Psychol.* **49**: 764–766.
44. Cowan N. 2011. The focus of attention as observed in visual working memory tasks: Making sense of competing claims. *Neuropsychologia* **49**: 1401–1406.
45. McCabe D.P. 2008. The role of covert retrieval in working memory span tasks: Evidence from delayed recall tests. *J. Mem. Lang.* **58**: 480–494.
46. Vergauwe E. & N. Langerock. 2017. Attentional refreshing of information in working memory: Increased immediate accessibility of just-refreshed representations. *J. Mem. Lang.* **96**: 23–35.
47. Camos V. & P. Barrouillet. 2014. Attentional and non-attentional systems in the maintenance of verbal information in working memory: the executive and phonological loops. *Front. Hum. Neurosci.* **8**: 900.
48. Portrat S. & B. Lemaire. 2014. Is Attentional Refreshing in Working Memory Sequential? A Computational Modeling Approach. *Cogn. Comput.* 1–13.
49. Lemaire B., A. Pageot, G. Plancher, *et al.* 2017. What is the time course of working memory attentional refreshing? *Psychon. Bull. Rev.* 1–16.
50. Wilken P. & W.J. Ma. 2004. A detection theory account of change detection. *J. Vis.* **4**: 11.
51. Prinzmetal W., H. Amiri, K. Allen, *et al.* 1998. Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *J. Exp. Psychol. Hum. Percept. Perform.* **24**: 261–282.
52. Souza A.S. & Z. Skóra. 2017. The interplay of language and visual perception in working memory. *Cognition* **166**: 277–297.
53. Camos V., P. Lagner & P. Barrouillet. 2009. Two maintenance mechanisms of verbal information in working memory. *J. Mem. Lang.* **61**: 457–469.
54. Raye C.L., M.K. Johnson, K.J. Mitchell, *et al.* 2007. Refreshing: A minimal executive function. *Cortex* **43**: 135–145.
55. Sandry J., J.D. Schwark & J. MacDonald. 2014. Flexibility within working memory and the focus of attention for sequential verbal information does not depend on active maintenance. *Mem. Cognit.* **42**: 1130–1142.
56. Vergauwe E. 2017. Comparing different instructed-refreshing schedules: Evidence for cumulative, forward-order refreshing of verbal lists? *Manuscr. Submitt. Publ.*
57. Hardman K.O., E. Vergauwe & T.J. Ricker. 2017. Categorical working memory representations are used in delayed estimation of continuous colors. *J. Exp. Psychol. Hum. Percept. Perform.* **43**: 30–54.
58. van den Berg R., E. Awh & W.J. Ma. 2014. Factorial comparison of working memory models. *Psychol. Rev.* **121**: 124–149.
59. Bae G.Y., M. Olkkonen, S.R. Allred, *et al.* 2015. Why some colors appear more memorable than others: A model combining categories and particulars in color working memory. *J. Exp. Psychol. Gen.* **144**: 744–763.
60. Oberauer K., C. Stoneking, D. Wabersich, *et al.* 2017. Hierarchical Bayesian measurement models for continuous reproduction of visual features from working memory. *J. Vis.* **17**: 11–11.

Figure Captions

Figure 1. Flow of events in the Refreshing conditions used in Experiment 1a (panel A), Experiment 1b (panel B), Experiment 2 (panel C), and Experiment 3a (panel D). Experiments 3b and 3c were similar to Experiment 3a with the following exceptions. First, mode of recall was typed. Second, before and after the sequence of refreshing instructions, a 0.5 s blank interval was inserted to match the timing implemented in Experiments 1 and 2. Displays are not drawn to scale. Specific details about the size of the stimuli in each experiment can be found in the Online Supplementary Materials.

Figure 2. Data of the continuous location reproduction task. Panel a shows recall error in the Refreshing condition (separately for 0-, 1-, and 2-Refreshing items) and in the no-cuing Baseline (none) of Experiments 1a and 1b. Error bars depict 95% within-subjects confidence intervals. Panel b shows group-level estimates of the probability of recalling the target according to the best-fitting mixture model applied to the data of the Refreshing condition. Error bars indicate the 95% highest-density interval of the posterior distribution of the parameter.

Figure 3. Data of the continuous color reproduction task used in Experiment 2. Panel a shows the mean recall error as function of refreshing frequency. Error bars depict 95% within-subjects confidence intervals. Panel b shows group-level estimates of the probability of recalling the target according to the best-fitting mixture model. Error bars indicate the 95% highest-density interval of the posterior distribution of the parameter.

Figure 4. Proportion of correct words recalled in the Refreshing condition (as a function of refreshing frequency: 0, 1, 2) and in the no-cue Baseline (none) in the three experimental versions that differed regarding recall mode (Oral vs. Typed) and articulatory suppression (AS). Exp. 3a = Oral – AS; Exp.3b = Typed – AS; Exp. 3c = Typed – no AS.

Figure 5. Posterior probability distribution of the refreshing frequency effect (2-Refreshing vs. 0-Refreshing) across the six experiments. The posterior indicates the range of credible values of a parameter given the data. The mean of the posterior is shown in each panel alongside the 95% highest density interval (HDI) of the distribution (colored bar underneath it). The top panel shows the refreshing effect on visual-spatial materials (Exp. 1a = grey line; Exp. 1b = blue line; Exp. 2 = black line). The second panel shows the refreshing effect on verbal materials (Exp. 3a = grey line; Exp. 3b = blue line; Exp. 3c = black line). The red dotted line indicates the value under the Null hypothesis.

Table 1

General Features of the Experiments Reported Herein.

Exp.	**Domain**	**Stimuli**	**Presentation**	**AS**	**Cue**	**Test**
1a	visuospatial	location	simultaneous	no	peripheral dot	continuous reproduction
1b	visuospatial	location	simultaneous	no	central color	continuous reproduction
2	visuospatial	colors	sequential	yes	arrow	continuous reproduction
3a	verbal	words	sequential	yes	arrow	oral recall
3b	verbal	words	sequential	yes	arrow	typed recall
3c	verbal	words	sequential	no	arrow	typed recall

Note. AS = articulatory suppression.

Table 2

WAIC for the Mixture Models Fitted to the Data of Experiments 1 and 2.

Exp.	Model	Parameters		WAIC_1	WAIC_2
		P^M	σ		
1a	M_1	Refreshing	Refreshing	93878.1	93879.3
	M_2	**Refreshing**	**---**	**93874.8**	**93875.9**
	M_3	---	Refreshing	93884.9	93886.1
1b	**M_1**	**Refreshing**	**Refreshing**	**68334.2**	**68335.5**
	M_2	Refreshing	---	68336.0	68337.2
	M_3	----	Refreshing	68342.1	68343.4
1(a+b)	M_1	Refreshing	Refreshing	162212.2	162214.8
	M_2	**Refreshing**	**---**	**162210.8**	**162213.2**
	M_3	---	Refreshing	162227.0	162229.5
2	M_1	Refreshing	Refreshing	126473.9	126474.3
	M_2	**Refreshing**	**---**	**126472.1**	**126472.5**
	M_3	---	Refreshing	126480.9	126481.3

Note. Lower values indicate better fit. The best model is printed in bold.

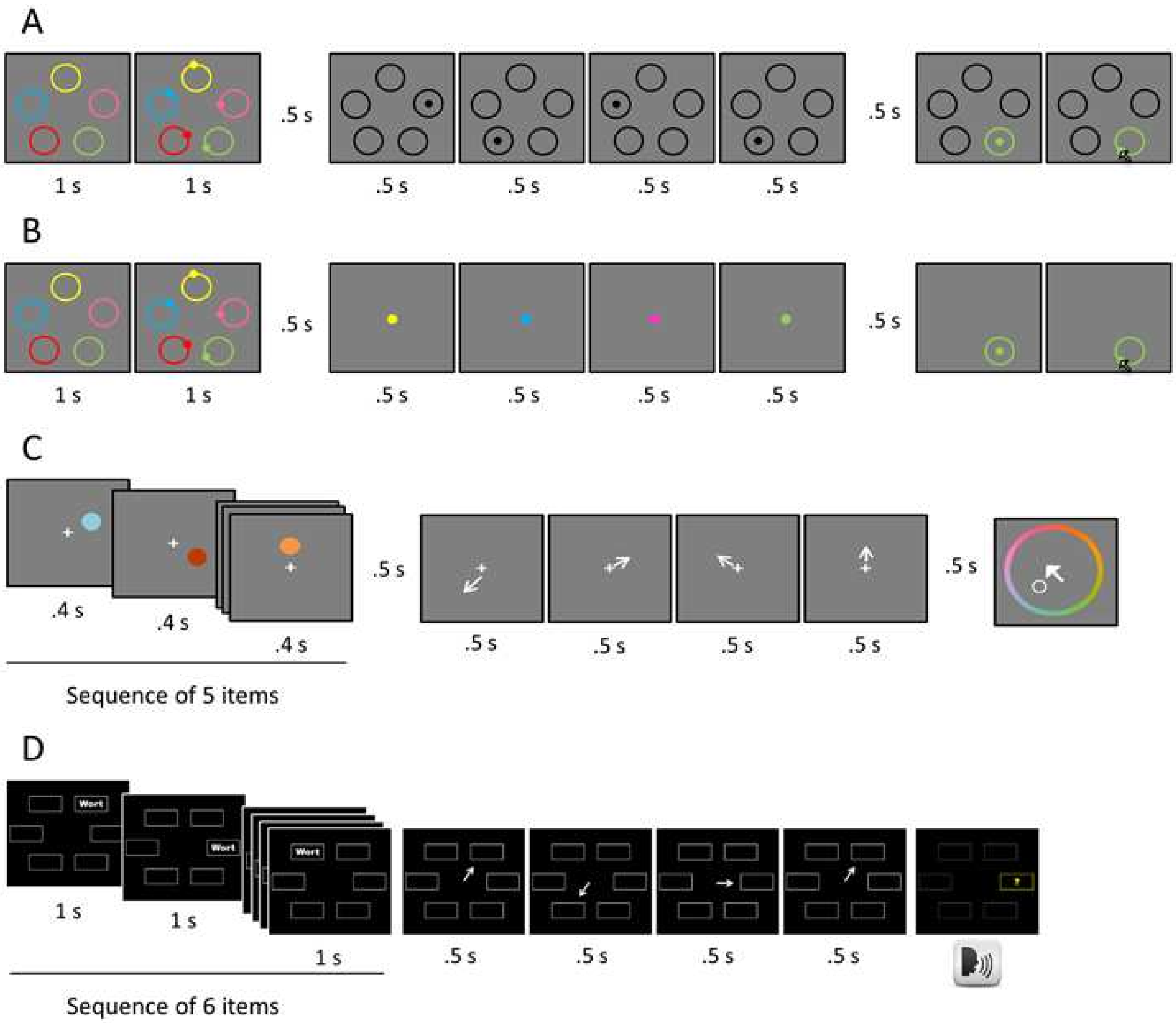

Figure 1. Flow of events in the Refreshing conditions used in Experiment 1a (panel A), Experiment 1b (panel B), Experiment 2 (panel C), and Experiment 3a (panel D). Experiments 3b and 3c were similar to Experiment 3a with the following exceptions. First, mode of recall was typed. Second, before and after the sequence of refreshing instructions, a 0.5 s blank interval was inserted to match the timing implemented in Experiments 1 and 2. Displays are not drawn to scale. Specific details about the size of the stimuli in each experiment can be found in the Online Supplementary Materials.

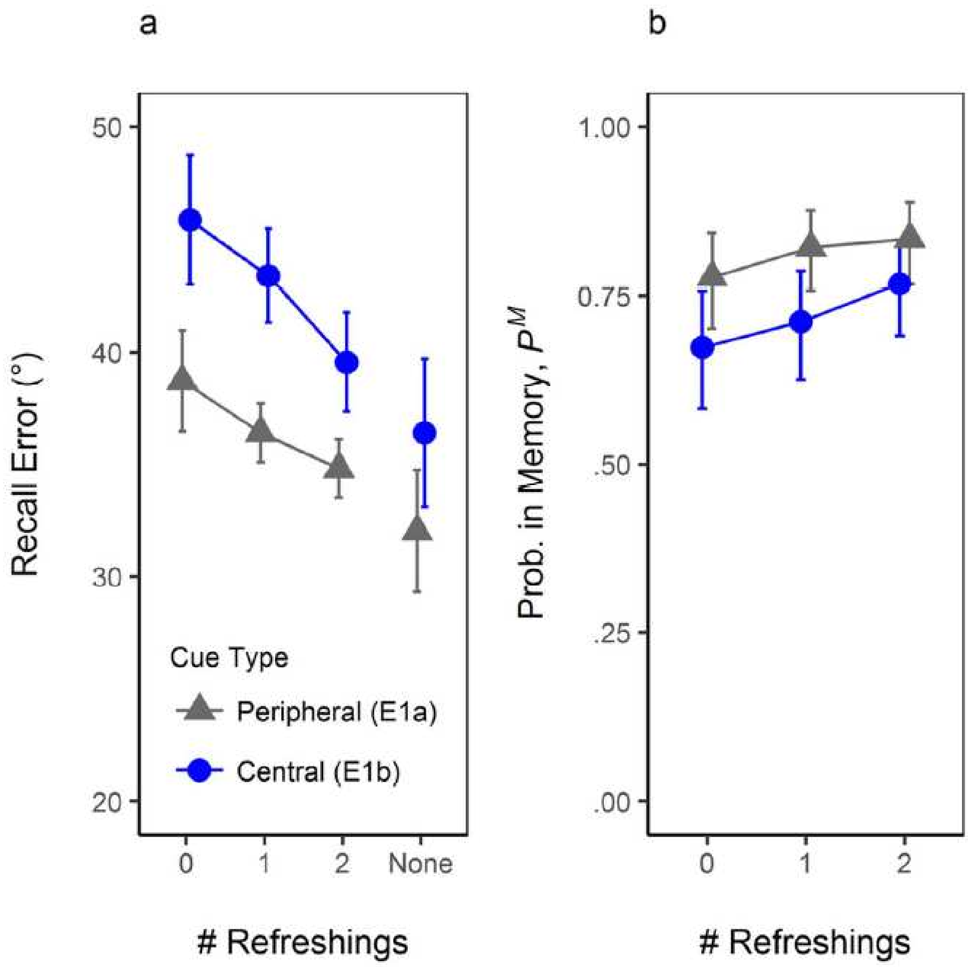

Figure 2. Data of the continuous location reproduction task. Panel a shows recall error in the Refreshing condition (separately for 0-, 1-, and 2-Refreshing items) and in the no-cuing Baseline (none) of Experiments 1a and 1b. Error bars depict 95% within-subjects confidence intervals. Panel b shows group-level estimates of the probability of recalling the target according to the best-fitting mixture model applied to the data of the Refreshing condition. Error bars indicate the 95% highest-density interval of the posterior distribution of the parameter.

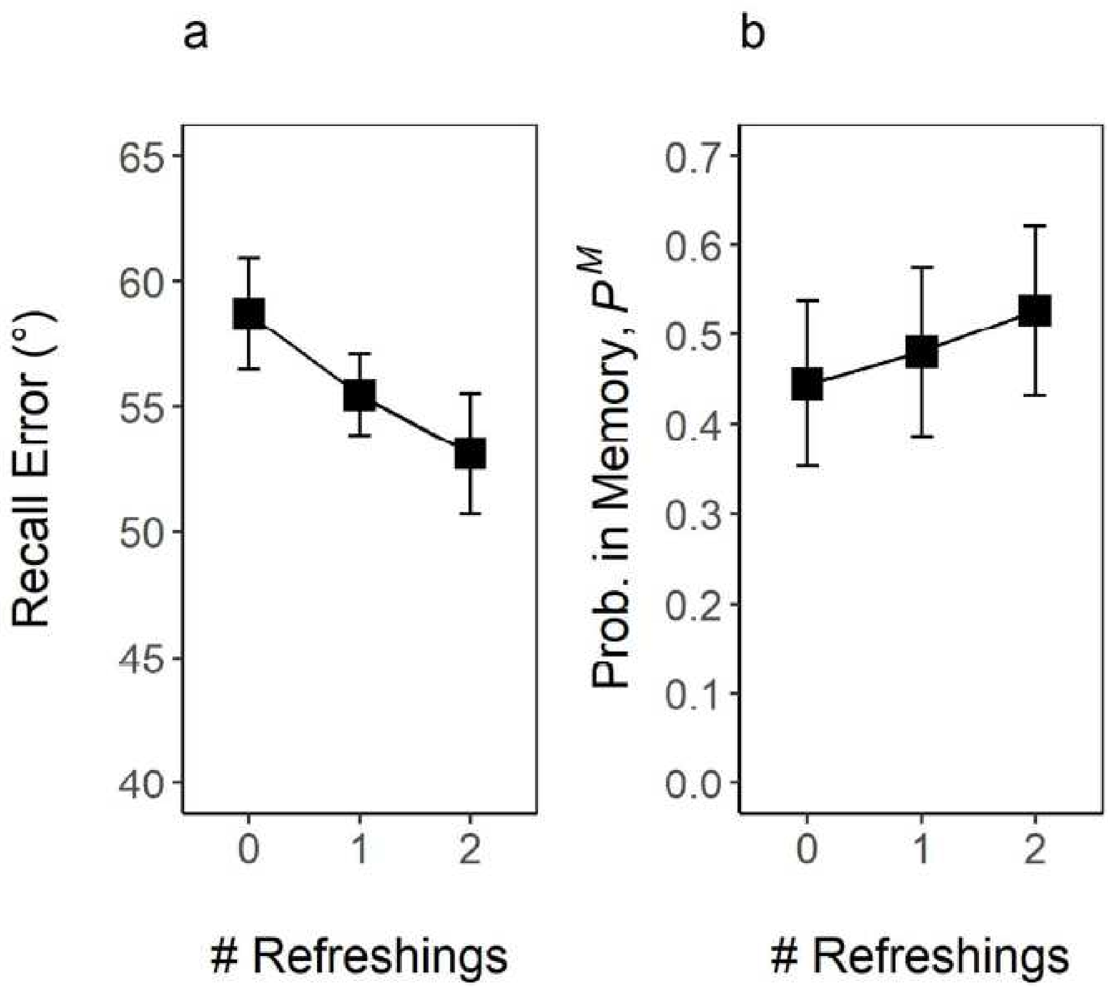

Figure 3. Data of the continuous color reproduction task used in Experiment 2. Panel a shows the mean recall error as function of refreshing frequency. Error bars depict 95% within-subjects confidence intervals. Panel b shows group-level estimates of the probability of recalling the target according to the best-fitting mixture model. Error bars indicate the 95% highest-density interval of the posterior distribution of the parameter.

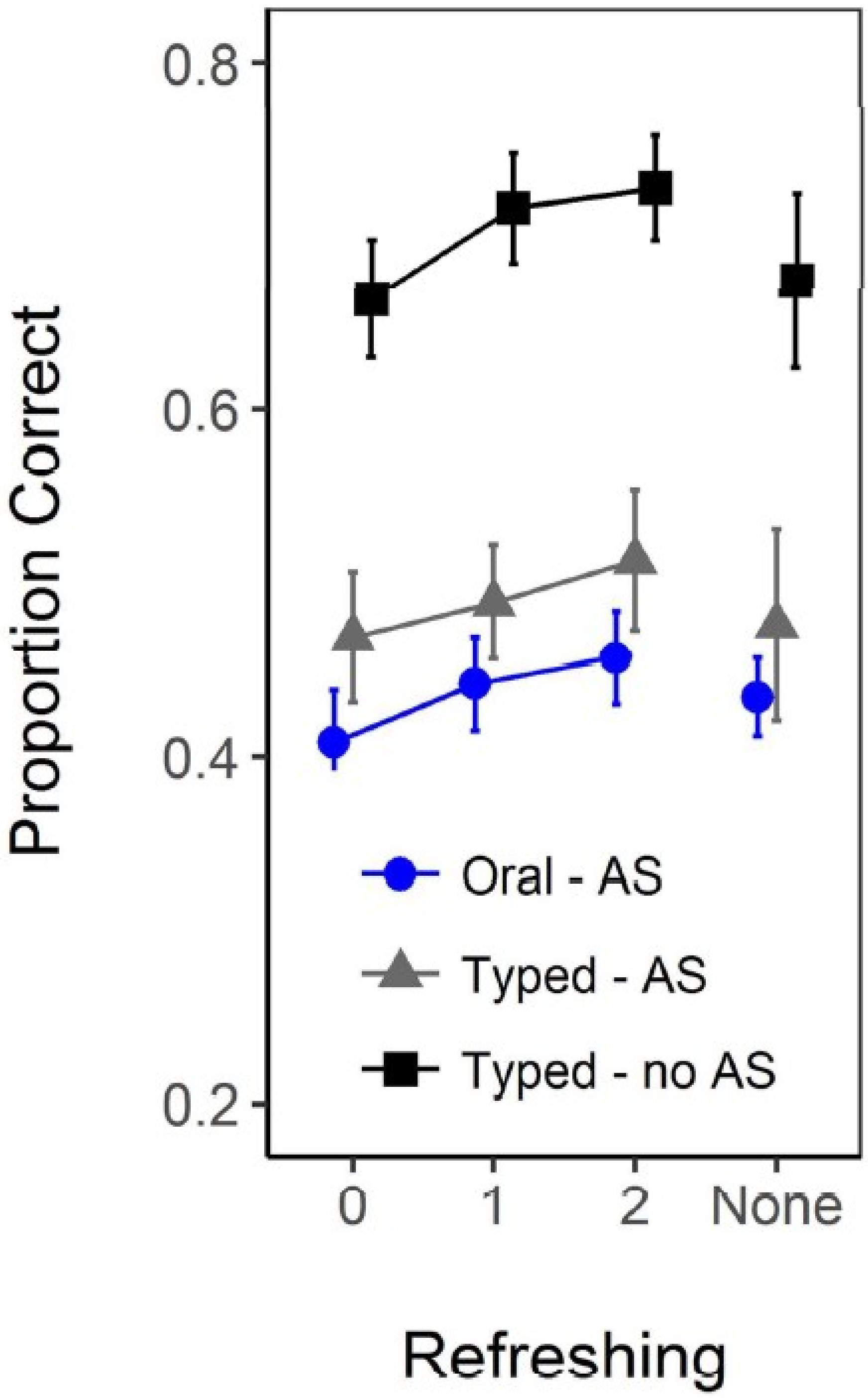

Figure 4. Proportion of correct words recalled in the Refreshing condition (as a function of refreshing frequency: 0, 1, 2) and in the no-cue Baseline (none) in the three experimental versions that differed regarding recall mode (Oral vs. Typed) and articulatory suppression (AS). Exp. 3a = Oral – AS; Exp.3b = Typed – AS; Exp. 3c = Typed – no AS.

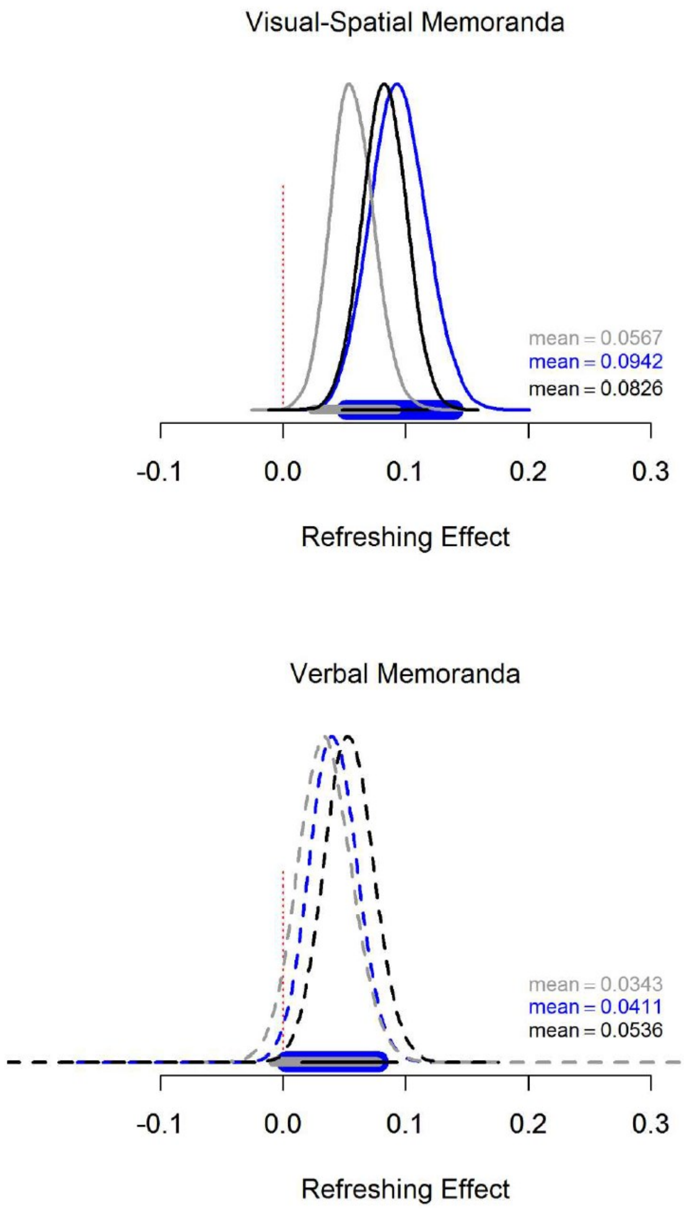

Figure 5. Posterior probability distribution of the refreshing frequency effect (2-Refreshing vs. 0-Refreshing) across the six experiments. The posterior indicates the range of credible values of a parameter given the data. The mean of the posterior is shown in each panel alongside the 95% highest density interval (HDI) of the distribution (colored bar underneath it). The top panel shows the refreshing effect on visual-spatial materials (Exp. 1a = grey line; Exp. 1b = blue line; Exp. 2 = black line). The second panel shows the refreshing effect on verbal materials (Exp. 3a = grey line; Exp. 3b = blue line; Exp. 3c = black line). The red dotted line indicates the value under the Null hypothesis.